AD________________

Award Number: W81XWH-12-1-0561

TITLE: Development of the Ovarian Cancer Cohort Consortium: Risk Factor Associations by Heterogeneity of Disease

PRINCIPAL INVESTIGATOR: Dr. Shelley Tworoger

CONTRACTING ORGANIZATION: Brigham and Women's Hospital, Inc., The

Boston, MA 02115-6110

REPORT DATE: october 2014

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE	**2. REPORT TYPE**	**3. DATES COVERED**
October 2014	Annual	30 Sep 2013 – 29 Sep 2014

4. TITLE AND SUBTITLE
Development of the Ovarian Cancer Cohort Consortium: Risk Factor Associations by Heterogeneity of Disease

5a. CONTRACT NUMBER

5b. GRANT NUMBER
W81XWH-12-1-0561

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Dr. Shelley Tworoger

E-Mail: nhsst@channing.harvard.edu

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Brigham and Women's Hospital, Inc., The
75 Francis Street
Boston, MA 02115-6110

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT
The objective of this application is to study the etiologic heterogeneity of ovarian cancer in multiple cohorts and to build the infrastructure of the Ovarian Cancer Cohort Consortium (OC3), an international consortium of cohort studies, to address scientific aims important for understanding ovarian cancer risk, early detection, and tumor heterogeneity that are only feasible in a consortium setting. Specifically we will examine associations of risk factors with invasive ovarian cancer, including (but not limited to) age, OCs, tubal ligation, parity, postmenopausal hormone use, family history of ovarian cancer, BMI, height, analgesic use, and lifetime ovulatory cycles, differ by histologic subtype, tumor dominance (as a surrogate for cell of origin), and tumor aggressiveness (tumors fatal within three years vs. all others). Then we will determine if risk prediction models for ovarian cancer can be improved by accounting for differential associations by cancer phenotype. In addition, the proposed efforts will create an infrastructure with a core dataset of important variables for ovarian cancer epidemiology that will be available for future efforts to study ovarian cancer risk, including projects that will use prospectively collected biological specimens. Currently, 23 cohorts have agreed to participate in the OC3. We have executed data use agreements between the Brigham and Women's Hospital (data coordinating center) with all studies. We have received data from 20 cohorts, with 3 cohorts actively preparing data. Data harmonization is complete for the cohorts for which we have received data. Analyses of primary ovarian cancer risk factors (e.g., oral contraceptive use, parity) by histology are complete and a manuscript is drafted and sent to co-authors. Preliminary analyses are on-going for the manuscript examining risk factors by tumor aggressiveness and for the development of a baseline risk prediction model. Several projects have been funded to utilize this resource and additional data collection and variable harmonization are underway.

15. SUBJECT TERMS
Ovarian Cancer, tumor heterogeneity, histology, cell of origin, tumor aggressiveness, risk prediction

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			USAMRMC
U	U	U	UU	31	**19b. TELEPHONE NUMBER** *(include area code)*

Table of Contents

Progress Report for OC110197
Sept. 2014

INTRODUCTION

The objective of this translational leverage award is to study the etiologic heterogeneity of ovarian cancer in multiple cohorts and to build the infrastructure of the Ovarian Cancer Cohort Consortium (OC3). The OC3 is an international consortium of cohort studies designed to address scientific aims important for understanding ovarian cancer risk, early detection, and tumor heterogeneity. The OC3 is part of the NCI Cohort Consortium, which is an extramural-intramural partnership to address the need for large-scale collaborations and provides the super-structure (but not funding) for managing the OC3. The OC3 currently has 23 participating, on-going cohort studies and we expect there to be over 8,000 ovarian cancer cases among more than 1.5 million women. The goals of the OC3 are to bring together cohorts with ovarian cancer endpoints for pooled projects, build a focused group of ovarian cancer researchers, and develop a comprehensive approach that integrates questionnaire and pathology data with biomarkers, genetics, and tissue. In addition to building the OC3 infrastructure, we propose to evaluate associations of ovarian cancer risk factors by different metrics of tumor heterogeneity. The first specific aim of this application is to examine whether associations of known and putative ovarian cancer risk factors, including (but not limited to) age, oral contraceptive use, tubal ligation, parity, postmenopausal hormone use, family history of ovarian cancer, body mass index, height, analgesic use, and lifetime ovulatory cycles, differ by (a) histologic subtype, (b) tumor dominance (as a surrogate for cell of origin), and (c) tumor aggressiveness (tumors fatal within three years vs. all others). We will use the information generated in the first aim to conduct the second specific aim, which is to develop ovarian cancer risk prediction models accounting for differential associations by cancer phenotype.

KEYWORDS
Ovarian Cancer, tumor heterogeneity, histology, cell of origin, tumor aggressiveness, risk prediction

OVERALL PROJECT SUMMARY

This grant began on September 30, 2012. Currently, 23 cohorts have agreed to participate in projects addressing the risk factor associations by tumor heterogeneity and to develop an improved risk prediction model for ovarian cancer. The tasks completed in the second year included: (1) finalization of data use agreements, (2) having cohorts prepare data and send to the Brigham and Women's Hospital (BWH) data coordinating center (DCC) for harmonization, (4) when available, conducting additional pathologic abstraction, (5) completing harmonization of core variables at the DCC, and (6) conducting statistical analyses for our aims.

A data dictionary was developed by the OC3 Steering Committee, and the data dictionary and a short questionnaire about the data collection and attributes were sent to all interested cohorts (see 2013 progress report). Only a subset of 10 cohorts have collected pathology reports – we currently are working with these cohorts to have personnel from the DCC to travel to the site and conduct the abstraction directly, work is completed for 2 cohorts, on-going from 3 others, and plans are being set up for the remaining sites. Due to the continuing difficulties in obtaining the appropriate permissions to access the pathology reports as well as pulling pathology reports out of long term storage, this work will continue into year 3.

In total, 23 cohorts from the US, Australia, Europe, and Asia have agreed to participate. For IRB purposes we require that each cohort establish a data use agreement (DUA) or provide a letter stating that the IRB does not require a DUA (if sending completely de-identified data). Three cohorts are primed at the BWH and therefore are covered under the primary IRB protocol and thus do not need a DUA. We have received executed DUAs between the BWH and 19 cohorts; 1 cohort did not require an agreement. We have received data from 20 cohorts. Details of the participating cohorts including sample sizes are presented in Table 1. Three cohorts have agreed to participate and are actively preparing data including the Adventist Health Study II, Melbourne Collaborative Cohort Study, and the Swedish Mammography Cohort. Delays in data receipt are due to a variety of factors including medical/maternity leaves and change in study PI; we hope to receive data by the end of 2014. We also have formally invited two large cohorts that have expressed possible interest in participating: the Women's Health Initiative in the US and the Million Women Study in the UK. Our policies and procedures are at our website: https://sites.google.com/a/channing.harvard.edu/oc3/.

Study name (ref)	Study code	Location	Enrollment period	Baseline cohort size*	Median participant age	End of follow-up	N, ovarian cases
NIH-AARP Diet and Health Study	AARP	U.S.	1995-1997	140,584	62	2006	761
Breast Cancer Detection Demonstration Follow-up Study	BCDDP	U.S.	1987-1989	36,080	61	1999	170
Breakthrough Generations Study	BGS	U.K.	2003-2011	101,905	48	2011	106
Canadian National Breast Screening Study	CANADA	Canada	1991-1999	2,765	58	2011	112
Campaign against Cancer and Stroke	CLUE II	U.S.	1989	5,673	47	2008	39
Cancer Prevention Study II Nutrition Cohort	CPS2	U.S.	1992-1993	65,975	62	2009	551
California Teachers Study	CTS	U.S.	1995-1999	43,810	50	2010	213
European Prospective Investigation into Cancer and Nutrition Study	EPIC	Europe / U.K.	1992-2000	264,461	51	2010	949
Iowa Women's Health Study	IOWA	U.S.	1986	30,716	61	2010	389
Multiethnic/Minority Cohort Study	MEC	U.S.	1993-1998	16,506	57	2011	107
Nurses' Health Study	NHS	U.S.	1980	94,868	46	2010	905
Nurses' Health Study II	NHS2	U.S.	1989	111,964	35	2011	279
NY University Women's Health Study	NYU	U.S.	1984-1991	12,440	49	2012	138
Netherlands Cohort Study on diet and cancer	NLCS	U.S.	1986	2,757	62	2009	448
Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial	PLCO	U.S.	1993-2002	60,266	62	2011	410
Singapore Chinese Health Study	SCHS	Singapore	1993-1999	31,962	56	2008	113
Sisters Study	SISTERS	U.S.	2003-2009	39,197	55	2012	40
VITamin D and OmegA-3 TriaL	VITAL	U.S.	2000-2002	28,331	60	2010	130
Women's Lifestyle and Health	WLHS	Sweden	1991-1992	49,218	40	2012	332
Women's Health Study	WHS	U.S.	1993-1996	33,584	53	2012	228
Total: 20 studies				1,173,062			6,420

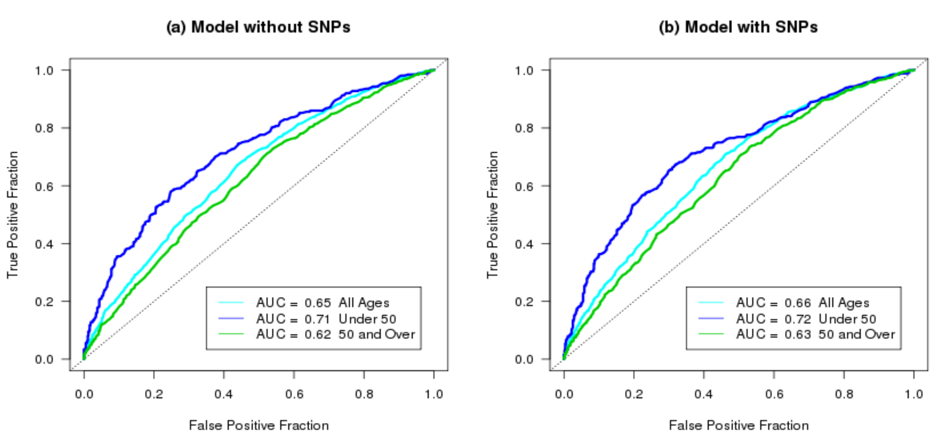
Table 1. Characteristics of 20 studies that have sent data to the data coordinating center for the OC3

*Eligible for inclusion in our analyses, including having a least one ovary and no baseline cancer
^Sent as a case-cohort study

Data harmonization for the key variables to be used in this analysis is complete for 20 cohorts from which we have received data. Specifically we have cleaned and harmonized the following variables: ovarian cancer diagnosis characteristics (date/age of diagnosis, date of death, type of tumor, morphology, histology, grade), study enrolment and follow-up data (date/age of enrolment, date/age of death, date/age of last follow-up), race, prior cancer diagnoses, family history of ovarian or breast cancers, menopausal status, postmenopausal hormone use (ever/never, duration, and type), use of oral contraceptives (ever/never, duration), tubal ligation, parity, hysterectomy status, oophorectomy status, age at menarche, age at menopause, smoking, height, body mass index (BMI), BMI at age 18, alcohol intake, endometriosis, other cancer diagnoses, and NSAID/aspirin use. Other variables collected through the OC3 will be cleaned at a later time when additional funding is obtained. We are now finalizing a detailed data dictionary (Appendix 1) explaining data decisions, categorizations, etc. to improve efficiency of future analyses.

In evaluating existing risk prediction models that could be used as the comparison model for the risk prediction aim, it became clear that existing studies have been based on small sample sizes or had limited validation. Therefore we decided that it was crucial to develop a "base" risk prediction model that could be used as a comparison model to evaluate improvements from multiple perspectives (e.g., incorporating tumor heterogeneity or genetics) with a large sample size and a statistically rigorous approach. To establish such a model, we are collaborating with the Ovarian Cancer Association Consortium (OCAC), which is a consortium of case-control studies, to develop and validate a risk prediction model. Model building in the OCAC is complete and the paper has been submitted to Cancer Research. The risk estimates were stratified by age at diagnosis (<50, ≥50) and parity (nulliparous, parous) as these are potential modifying factors. We used generalized additive models (GAMs) with random effects for study site, fixed effects for categorical variables and SNPs, and smooth non-parametric functions for continuous variables. All risk factors, with the exception of age, had some missing data; 80% of the participants had at least one risk factor with missing values. Rather than limit the analysis to the participants with complete data or drop risk factors from the model, we imputed values of each of the missing risk factors as a function of other risk covariates as well as education level, smoking status, and alcohol use (Appendix 2). To do this, we developed a Bayesian graphical model that provided a coherent sequence of conditional models for the risk factors and indicators of whether they are missing. Using the observed and imputed data, we developed a multivariate hierarchical logistic regression model for predicting case-control status as a function of the epidemiological and genetic risk factors using Markov Chain Monte Carlo (MCMC); the multiple imputations through MCMC provide valid confidence intervals for statistical inference by addressing uncertainty in the missing values and reduce bias induced by complete case analyses when data are not missing at random. Models with and without the 11 SNPs were fit to the training data (random sample of 80%) and used to predict case-control status on the validation data (remaining 20%). Point estimates of log odds ratios were estimated by the median of the samples from the posterior distribution of each of the parameters; Bayesian 95% confidence intervals (CI) were obtained by taking the 2.5th percentile and 97.5th percentile of the estimated posterior distribution for each parameter. Model parameters are shown in Appendix 2. Performance of the model is shown in Figure 1 below.

Figure 1. Receiver operating characteristic curve for models (a) without and (b) with SNPs

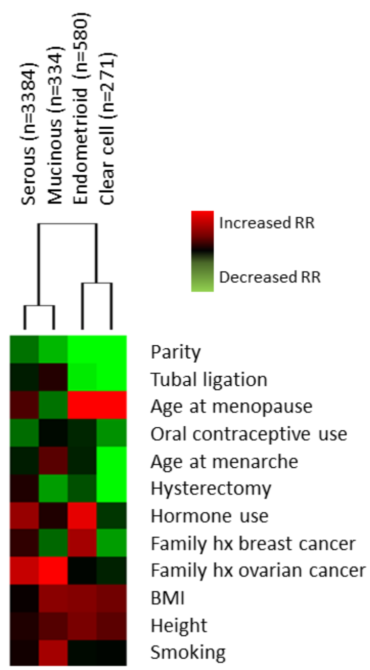


Overall the model fit better for younger women and the addition of SNPs added little to the predictive value of the model. We are now using the parameters derived from the OCAC data as the prior distribution to refit the model in the OC3 cohorts. We are using US-based cohorts only (to more easily obtain information for calculating baseline risk of disease, e.g., bilateral oophorectomy rates, cancer rates), and are holding out 2 cohorts (CPS2 and NHS2) for validation. Additional model development will be conducted on a random 80% of participants from the other cohorts (AARP, BCDDP, CLUEII, CTS, IOWA, MEC, NHS, NYU, PLCO, VITAL, WHS), with an initial validation step on the remaining 20% of women in these cohorts. Thus far in the OC3, we have developed a model for baseline risk of ovarian cancer that accounts for the competing risks 'mortality' and 'diagnosis of another cancer' and that adjusts for bilateral oophorectomy (BSO) status, birth cohort and age. We used SEER mortality and incidence tables and associated error estimates to construct a prior distribution for the baseline incidence and mortality curves. We conducted a meta-analysis of ovarian cancer risk reduction and formed a prior distribution for the relative risk of ovarian cancer associated with BSO on basis of this analysis. Finally, we constructed a prior distribution on age-specific incidence of BSO on basis of an analysis of NHANES study data. These prior data have been saved as data structures in the R statistical environment. We developed a Markov Chain Monte Carlo (MCMC) algorithm, implemented in the R programing language using the JAGS package, to fit the model to data and have tested this computer code on the Harvard system using simulated data. In doing so, we have verified that this code generates estimates of the BSO incidence and relative risk parameters that are very close to the 'true' values used to simulate the data, giving us confidence that the algorithm is correct. We are currently carrying out minor recodes to the OC3 incidence and BSO data and making minor adjustments to the baseline model and MCMC algorithm to facilitate imputation of missing data. Once complete, we will fit the baseline model to the OC3 Phase I training set and evaluate its incidence predictions using the Phase I evaluation set. We will evaluate several versions of the baseline model that result from relaxing or tightening the prior distributions assumed for the model's BSO parameters. Once we are confident that the baseline model is correctly formulated and coded, we will proceed with the addition of risk factors and the model infrastructure necessary for imputing their missing data. This process will borrow heavily from existing infrastructure developed in context of the OCAC/PostGWAS relative risk modeling exercise. Because of the complexities of the data structure for this analysis and incorporating substantial input on the OCAC portion of the analysis from co-investigators, this project will be completed in the next 6 months. This unique collaboration will provide a resource for all future work on ovarian cancer risk prediction, including the incorporation of differential associations by histology.

We have developed SAS macros for conducting analyses in a standardized manner, including a macro to meta-analyze results for a particular exposure across studies, one to conduct a pooled analysis, and macros to assess risk factor association heterogeneity by tumor subtype. We have completed the analysis for examination of ovarian cancer risk factors by histology and a draft manuscript has been circulated to co-authors. Also, given the complexity of the data harmonization, we are doing an independent code review of the programs to ensure the highest quality analysis. The analysis includes 1,233,757 women and 4,569 cases with known histology of serous/poorly differentiated (n=3,384), endometrioid (n=580), mucinous (n=334), and clear cell (n=271) types. We calculated hazard ratios (HR) and 95% confidence intervals (95%CI) using competing risks Cox proportional hazards models to evaluate the association between key ovarian cancer risk factors and risk of ovarian cancer by histologic type. Follow-up time was defined as time between study entry and 1) date of ovarian cancer diagnosis, 2) date of death, or 3) end of follow-up, whichever occurred first. Two methods were used to estimate associations in this analysis. First, we pooled data from all cohorts. Models were stratified on age at study entry, year of birth, and cohort to account for potential differences in baseline hazards by these factors. All models were adjusted for number of children and duration of OC use, unless the exposure of interest was collinear with these factors. Analysis of hysterectomy was additionally adjusted for HT use. Statistical heterogeneity of associations across subtypes was assessed via a likelihood ratio test comparing a model allowing the association for the risk factor of interest to vary by histology versus one not allowing the association to vary. Second, we used random effects meta-analysis to combine cohort-specific estimates, which were produced as described above (without stratification by cohort).

Results for this analysis are shown in Appendix 3; notably, most putative ovarian cancer risk factor associations differed by histology. Compared to nulliparous women, parous women had a reduced risk of all ovarian cancer subtypes; however there was substantial heterogeneity in the amount of risk reduction across subtypes (p-het=3.49^{-08}). The strongest reduction was observed for clear cell cancers (RR: 0.34; 95% CI: 0.25-0.45), while serous cancers had the lowest risk reduction (RR: 0.78; 95% CI: 0.70-0.87). Similar patterns were observed among parous women for number of children. Age at menopause associations differed by subtype (p-het=0.01), with associations for endometrioid (RR: 1.20; 95% CI: 1.05-1.37) and clear cell carcinomas (RR: 1.37; 95% CI: 1.14-1.65). Ever use of menopausal hormone therapy was associated with increased risk of serous (RR: 1.39; 95% CI: 1.28-1.51) and endometrioid (RR: 1.65; 95% CI: 1.33-2.05) carcinomas, but not with mucinous or clear cell subtypes (p-het=0.01). Tubal ligation was associated with reduced risk of endometrioid (RR: 0.60 95% CI: 0.41-0.88) and clear cell carcinomas (RR: 0.35; 95% CI: 0.18-0.69), but not the other types (p-het=0.003), while hysterectomy was associated with mucinous (RR: 0.68; 95% CI: 0.48-0.97) and clear cell (RR: 0.57; 95% CI: 0.37-0.89) carcinomas only (p-het=0.01). No heterogeneity was observed for duration of breastfeeding, age at menarche, or OC use (p-het≥0.13), although for ever/never use, the association was significant only for serous (RR: 0.79; 95% CI: 0.73-0.85) and clear cell cancers (RR: 0.73; 95% CI: 0.55-0.96). A first degree family history of breast cancer was associated with endometrioid carcinomas (RR: 1.43; 95% CI: 1.11-1.84) and marginally with serous carcinomas (RR: 1.11; 95% CI: 1.00-1.24) with significant heterogeneity across subtypes (p-het=0.02). Interestingly, first degree family history of ovarian cancer was significantly associated only with serous carcinomas (RR: 1.54; 95% CI: 1.25-1.90), although there was no statistical heterogeneity (p-het=0.39), likely due to the rarity of the exposure. Body mass index was most strongly associated with non-serous subtypes, most strongly with endometrioid carcinomas. While none of the individual associations reached statistical significance, there was significant heterogeneity across the subtypes (p-het=0.04). Ever smoking was associated with mucinous carcinomas (RR=1.43; 95% CI: 1.14-1.80), but not with the other subtypes. The association with mucinous carcinomas increased with increasing number of pack years. Conversely, in the highest pack-year categories, we observed a significant reduction in risk of clear-cell carcinomas compared to never smokers (RR: 0.46; 95% CI 0.23-0.95 for >20-35 pack years and RR: 0.41; 95% CI 0.18-0.93 for >35 pack-years). No heterogeneity by subtype was observed for height. To agnostically evaluate patterns of risk factor associations in histologic subtypes of ovarian cancer, we performed unsupervised hierarchical clustering including all risk factors evaluated and the four histologic subtypes as endpoints (Figure 2). Across all exposures, each subtype had unique patterns of risk factor associations. Generally, most risk factors had their strongest association with non-serous cancers. Unsupervised clustering divided the histologic subtypes into two groups based on the similarity of risk factor associations, with serous and mucinous carcinomas in one group and endometrioid and clear cell carcinomas in the other group.

Figure 1: Unsupervised hierarchical clustering of ovarian cancer subtypes by risk factor associations



We also have completed preliminary analyses examining risk factor associations by tumor aggressiveness. Results are shown in Appendix 4. We defined tumor aggressiveness as rapidly fatal (death within 3 years) and less aggressive (all others). Ovarian cancer characteristics and diagnosis dates were abstracted from pathology reports or cancer registries; registries were used to ascertain vital status and date of death. We are awaiting death information specifically on ovarian cancer cases from 3 studies before we can finalize the analysis. As with the histology analysis we used competing risks Cox proportional hazards models and likelihood ratio tests. Among 4,066 cases with known vital status and the potential for at least 3 years of post-diagnosis follow-up, 2104 (51.7%; median survival=1yr) were rapidly fatal and 1,962 (48.3%; median survival=18yr) were less aggressive. 66% of rapidly fatal and 55% of less aggressive cases were of serous histology. Stronger associations were observed for less aggressive than for rapidly fatal disease for tubal ligation (RR, yes vs. no=0.69 vs. 1.06, respectively; p-het=0.003), parity (RR, per child=0.88 vs. 0.92; p-het=0.02), pack years of smoking (RR per 20 pack-yr=0.94 vs. 1.06; p-het=0.01), and suggestively for family history of breast cancer (RR, yes vs. no=1.19 vs. 1.01, p-het=0.10). Conversely, women with a BMI>35 vs. >22-25 kg/m^2 were at higher risk of rapidly fatal disease (RR=1.43), but not less aggressive disease (RR=1.01; p-het=0.07). Interestingly, we conducted these analyses stratified by histology (serous/poorly differentiated and endometrioid/clear cell). For example, tubal ligation was associated with a reduced risk of less aggressive disease for both categories of histology, but parity and family history of breast cancer were only associated with less aggressive disease of the endometrioid and clear cell types. Conversely, BMI and smoking were associated with an increased risk of rapidly fatal disease, primarily for the serous type. Analyses for this aim will be complete by the end of 2014 and the paper will be drafted and distributed to co-authors in early 2015.

With respect to the OC3 structure, we continue to have monthly conference calls run by the PI with the Steering Committee (Table 3). The calls focus on discussing on-going and future collaborations or projects, and vetting preliminary results. The PI also meets weekly with Dr. Elizabeth Poole (a junior faculty member working on the project) and the OC3 programmer. The OC3 has had three in-person meetings since the grant started, including at the 2013 Annual NCI Cohort Consortium Meeting. Our next in-person meeting is in December 2014 at the upcoming Cohort Consortium annual meeting. We chose these meeting times because many investigators attend these associated meetings so we have very good attendance. We also have developed a website for the OC3 to communicate our goals, guidelines for participation, and in the future, interesting findings from the study (see https://sites.google.com/a/channing.harvard.edu/oc3/?pli=1).

Seven projects have been proposed using the infrastructure of the OC3 in addition to the aims funded by this grant (Appendix 5). These projects have or will obtain separate funding to support the analyses needed for that project or any additional data collection. Data collection is on-going for 2 studies with Dr. Rudolph Kaaks and Renee Fortner to examine existing blood levels of androgens and IGFs in relation to risk, particularly by histology, and data cleaning is on-going for two studies with Drs. Nicolas Wentzensen and Britton Trabert, to examine the relationship of NSAIDs and endometriosis with ovarian cancer risk. We also are participating in the NCI-initiated OncoArray project, which includes a GWAS backbone plus specific SNP content for several cancer types, including ovarian cancer. Eight OC3 cohorts with appropriate biospecimens plan to participate and send samples using a nested case-control study design (Table 2). The OC3 will coordinate data management with the OCAC (who is managing the overall ovarian OncoArray project) and will add the genotyping data to the OC3 database after the initial analyses are complete within the full OncoArray consortium (genotyping and data cleaning expected to be complete at the end of 2014).

Study	Cases with DNA source	DNA source	Control selection	Status
SMC	105	blood	age	Genotyping complete; QC process on-going
Sisters	150	blood	age	Genotyping complete; QC process on-going
MEC	150	blood/ buccal	site, individual age	Genotyping complete; QC process on-going

PLCO	346	blood/ buccal	freq matched 5-year	Genotyping complete; QC process on-going
NHS / NHSII	380	blood/ buccal	age, DNA source, menopause	Genotyping complete; QC process on-going
VITAL	140	buccal	age	DNA extraction on-going
CLUE	100	blood	age	DNA extraction on-going
EPIC	890	blood	age, menopause	Genotyping complete; QC process on-going
TOTAL: 9studies	**2,261**			

KEY RESEARCH ACCOMPLISHMENTS

Below is a list of key research accomplishments in the second year of this award.

- Recruited 23 cohorts to the OC3, received data for 19 studies, invited 2 additional large studies
- Harmonized core dataset for 19 studies
- Enhanced and tested SAS macros for data analyses testing associations by tumor subtype
- Conducted analyses assessing risk factor associations by histologic subtype, drafted manuscript and sent to co-authors for review
- Submitted manuscript on the initial development of the risk prediction model for ovarian cancer in the OCAC
- Set up data structures and programs to run and validate a risk prediction model for ovarian cancer overall in the OC3, including incorporating baseline risk in a prospective context
- Ran preliminary analyses for aim evaluating associations by tumor aggressiveness
- Presented preliminary results on the histology analysis at the 2014 Society for Epidemiologic Research meeting and for the tumor aggressiveness analysis at the 2014 Ovarian Cancer Research Symposium sponsored by AACR and the Marsha Rivkin Center for Ovarian Cancer Research
- Continued coding tumor dominance in studies with pathology reports
- Continued participation by OC3 cohorts and investigators in monthly steering committee meetings and bi-annual in person meetings
- Multiple projects proposed to use OC3 resource with additional funding obtained by several investigators and more projects being proposed

CONCLUSION

We are actively developing the OC3 infrastructure by pooling existing cohort data to better elucidate the biology of ovarian cancer. Scientifically, we have or will evaluate whether associations for putative ovarian cancer risk factors differ by tumor subtypes (histology, cell of origin, aggressiveness), as well as develop risk prediction models based on differing risks across subtypes. Further, we are working to develop a “base” risk prediction model that can be used as a comparison for assessing improvement in future work. This will be beneficial to the entire ovarian cancer research community. Importantly in our initial work we observed that most established or putative ovarian cancer risk factors showed heterogeneity across histologic subtypes and all subtypes had unique patterns of risk factor associations. Endometrioid and clear cell tumors had the strongest associations for many risk factors, and relatively few associations were observed for serous tumors, which are the most common tumor type. This suggests that risk prediction models of ovarian cancer overall will perform worse for serous tumors than for other types. Further, our initial results comparing risk factors for rapidly fatal versus less aggressive disease suggests that this construct adds biologic information beyond that of histology. Overall, our results have several important implications for etiology and prevention of ovarian cancers. The substantial heterogeneity of individual risk factor associations across ovarian cancer subtypes supports the notion that the subtypes are indeed different diseases and that we may need to consider multiple tumor characterizations to adequately stratify tumors. This underscores the importance of evaluating risk factor and biomarkers associations in consortium settings where there is adequate sample size to provide power to assess associations for the more rare tumor types. The research also suggests that we need to identify new epidemiologic risk factors for serous tumors as the traditional factors are generally most strongly related to endemetrioid and clear cell tumors. Given the higher incidence of serous cancer and its poor survival rates, this is a critical area of future research.

This systematic approach to address ovarian cancer heterogeneity in a large consortial effort will set new standards for evaluating ovarian cancer risk factors and biomarkers and thereby impact understanding of ovarian cancer etiology beyond the work conducted in OC3. Further, the classification of ovarian cancers by histology, cell of origin, and aggressiveness in a large set of cohort studies sets an important harmonized framework for future risk and biomarker studies of ovarian cancer. This project makes extremely efficient use of existing data to produce new information that would otherwise not be available because of limited statistical power within individual studies as well as provide a basis for future consortial studies in the OC3.

PUBLICATIONS, ABSTRACTS, AND PRESENTATIONS

No publications at this time.
Two abstracts were accepted as presentations (presenter is bolded):

1. **Elizabeth M. Poole**, Alan A. Arslan, Lesley M. Butler, James V. Lacey, Jr., I-Min Lee, Alpa V. Patel, Kim Robien, Dale P. Sandler, Leo J. Schouten, V. Wendy Setiawan, Kala Visvanathan, Elisabete Weiderpass, Emily White, Nicolas Wentzensen, Shelley S. Tworoger. Ovarian cancer risk factors by histologic type in the Ovarian Cancer Cohort Consortium (OC3). Presented at the 2014 Annual Meeting of the Society for Epidemiologic Research, June 2014, Seattle, WA.
2. **Shelley S. Tworoger**, Elizabeth M. Poole, Alan A. Arslan, Lesley M. Butler, Victoria Kirsh, James V. Lacey, Jr., I-Min Lee, Alpa V. Patel, Kim Robien, Thomas Rohan, Dale P. Sandler, Leo J. Schouten, V. Wendy Setiawan, Kala Visvanathan, Elisabete Weiderpass, Emily White, Nicolas Wentzensen. Ovarian cancer risk factor associations by tumor aggressiveness in the Ovarian Cancer Cohort Consortium (OC3). Presented at the 10th Biennial Ovarian Cancer Research Symposium sponsored by AACR and the Marsha Rivkin Center for Ovarian Cancer Research, September 2014, Seattle, WA.

INVENTIONS, PATENTS, AND LICENCES

None.

REPORTABLE OUTCOMES

The primary reportable outcome is the development of the OC3 database, which contains data on ovarian cancer risk factors and outcomes from 20 cohort studies and by the end of 2014 will contain data from 3 more studies. This resource can be used for the analyses proposed in this grant as well as other analyses (see Appendix 5 for additional projects).

OTHER ACHIEVEMENTS

None.

REFERENCES

None.

APPENDICES

Appendix 1: Detailed data dictionary and coding decisions for harmonized variables in the OC3
Appendix 2: Tables from the risk model project with the OCAC describing imputation variables for missing data and the log odds ratio estimates from the prediction model
Appendix 3: Results for ovarian cancer risk factor associations by histology
Appendix 4: Preliminary results for ovarian cancer risk factor associations by tumor aggressiveness
Appendix 5: Additional proposed projects using the OC3 infrastructure and the status

Appendix 1. Detailed data dictionary and coding decisions for harmonized variables in the OC3.

Variable Name	Description	Coding	Comment
NEWID	Unique ID for each study participant	use sequential numbers within each study	oc3 own coding as some studies provide character id and some numeric
CASE CHARACTERISTICS			
INVASIVE	Ovarian cancer invasive case identifier	1=yes, invasive; 0=no, non-invasive ovarian case; missing for non case	derived from ovca and behavior to only include invasive epithelial and peritoneal case
	AARP, use BEHAVIOR to derive invasive	if behavior=3, malignant, primary site, then invasive=1; else if behavior=2, carcinoma in situ, then invasive=0; the rest set to missing	
	CLUE, use behavior to derive invasive	if behavior=2 and ovarycan=1 then invasive=1; else if behavior=2 and ovarycan=1 then invasive=0; the rest set to missing	
OVARYCAN	Ovarian caner case identifier	1=yes, epithelial or peritonial; 0=no	binary. Here 'yes' means it is either epithelial or peritoneal ovarian cancer case; '0' means non ovarian cancer case, as we deleted all other ovarian cancer cases.
	AARP doesn't have this variable. So derived from behavior	if behavior=3 or 2, malignant, primary site or carcinoma in situ, then ovarycan=1; the rest set to 0	
OVCA	Ovarian cancer case identifier	1=epithelial ovarian cancer case; 2=peritoneal case; missing if non-case	delete non epithelial or peritoneal ov cases, i.e., 3=fallopian tube case; 4=germ cell case; 5=sex cord case; missing if non-case
	AARP doesn't have this variable. So derived from behavior	if behavior=3 or 2, malignant, primary site or carcinoma in situ, then ovca=1; the rest set to 0	for AARP, ovarycan=ovary

	BCDDP & CANANDA didn't have this variable. So derived from ovarycan	if ovarycan=1 then ovca=ovarycan; else ovca=.	ovarycan is provided, 0=no, 1=yes. Presumably all epithelial ov cases
	AARP has different categories	2=carcinoma in situ; 3=malignant, primary site	assign 3 as 2; 2 as 1; change from character to numeric
QXYEAR	Year of baseline questionnaire return	Year of baseline questionnaire return as YYYY	for NYU, derived from birth year and QXAGE
QXMONTH	Month of baseline questionnaire return	Month of baseline questionnaire return as MM	PLCO & VITAL & NYU didn't provide
QXAGE	Age at baseline questionnaire return	Age at baseline questionnaire return as YY; . if unknown	for NYU, derived from birth year and AGE
DEATHYEAR	Year of death	Year of death as YYYY; missing if not dead by end of follow-up	for NYU, derived from birth year and DEATHAGE
DEATHMONTH	Month of death	Month of death as MM; missing if not dead by end of follow-up; . if unknown	PLCO & VITAL & NYU didn't provide
DEATHAGE	Age at death	Age at death as YY; missing if not dead by end of follow-up; . if unknown	
DXYEAR	Year of ovarian cancer diagnosis	Year of diagnosis as YYYY; missing if non-case	for NYU, derived from birth year and DXAGE
DXMONTH	Month of ovarian cancer diagnosis	Month of diagnosis as MM; missing if non-case	PLCO & VITAL & NYU didn't provide
DXAGE	Age at ovarian cancer diagnosis	Age at diagnosis as YY (in years); missing if non-case	the original was called agedx in months
LASTYEAR	Year of last follow-up	Year of last follow-up as YYYY	for NYU, derived from birth year and LASTAGE
LASTMONTH	Month of last follow-up	Month of last follow-up as MM	PLCO & VITAL & NYU didn't provide
LASTAGE	Age at last follow-up	Age at last follow-up as YY; . if unknown	
OVTIME	derived follow-up time	follow-up time in years	It is calculated using the minimum non-missing of last follow-up, ovarian cancer diagnosis, and death minus baseline questionnaire return (in number of years). The exception is NYU which provided derived qxage, dxage, lastage, and deathage. IOWA has provided this variable.
ENTRYAGE	derived age at questionnaire return to study	age in years (=qxage in years)	for competing risk analyses, we used Anderson-gill model: (entryage, eventage) in phreg.

EVENTAGE	derived age when event (dx ovarian cancer, last follow up, or death)	= MIN(agedx, lastage, deathage)	For case-cohort studies such as Canada & Netherlands, the entryage will be the eventage minus .001, i.e., almost no fu time, for case outside cohort.
DXTIME	derived dx time in months	= dxyear*12+dxmonth	
QXTIME	derived qx time in months	= qxyear*12+qxmonth	
LASTTIME	derived last follow up time in months	= last year*12+lastmonth	
DEATHTIME	derived death time in months	= death*12+deathmonth	
DEMOGRAPHIC DATA			
BIRTHYEAR	Year of birth	Year of birth as YYYY	Breakthrough only provided 5-year age group, so set medium
BIRTHMONTH	Month of birth	Month of birth from 1 to 12	PLCO & VITAL & NYU didn't provide
RACE	Race	1=White; 2=Black; 3=Asian/Pacific Islander; 4=Native American; 5=Other; 9=Unknown	
	AARP	1=non-Hispanic white; 2=non-hispanic black; 3=hispanic; 4=Asian; 5=pacific islander; 6=American Indian/Alaska native; 9=unknown	assign AARP 1 to 1; 2 to 2; 3 to 5; 4 & 5 to 3; 6 to 4; 9 to 9
HISPANIC	Ethnicity	1=Hispanic; 0=Non-Hispanic	
	AARP	N/A	derived from race, set AARP race=3 as hispanic=1; the rest as 0
EDUCTION	Highest level of education	1=did not finish high school; 2=high school; 3=some college; 4=completed college; 5=post graduate; 9=unknown/missing	
ANTHROPOMETRIC VARIABLES			
WEIGHT18	weight at age 18/20/21 in pounds	set missing if values <50 lb. or >600 lb.	BCDDP & NYU didn't provide this information. PLCO provided weight at age 20; Clue provided weight at age 21.
BMI18	BMI at age 18/20/21 in kg/m^2	set missing if values <13 or >60	derived from WEIGHT18 and HEIGHT

HEIGHT	Height in inches	set missing if values <48 inch or >84 inch	
BMI	body mass index	continuous, set missing if values <14 or >60	
MENARCHE AND MENOPAUSE VARIABLES			
AGEMENARCHE	Age in years when menstrual periods began	.=unknown, or values <5 or >30 years old	create continuous based on midpoint of each category if continuous is not available. CPS2, BCDDP and Nurses have continuous age at menarche.
	PLCO only has categorical AGEMENARCHE	1=<10; 2=10-11; 3=12-13; 4=14-15; 5=16+; 99=unknown;	assign values 9.0, 10.5, 12.5, 14.5, 16.5 and missing for categories 1, 2, 3, 4, 5 and missing respectively
	AARP only has categorical AGEMENARCHE	1=<=10; 2=11-12; 3=13-14; 4=>=15; 9=unknown	assign values 9.5, 11.5, 13.5, 16, and missing for categories 1, 2, 3, 4, and 9 respectively
	VITAL only has categorical AGEMENARCHE	1=<=11; 2=12; 3=13; 4=14; 5=15; 6=16+; 7=never had a period	assign values 11, 12, 13, 14,15, 16 respectively; and missing for the rest
	CTS only has categorical AGEMENARCHE	1=NOFMP; 2=<10; 3=10; 4=11; 5=12; 6=13; 7=14; 8=15; 9=1'; 10=17+	assign values 0, 9, 10, 11,12,13,14,15,16, 17 respectively; and missing for the rest
MENO	Menopause status	1=post; 0=pre; .=unknown; .=dubious menopause	AARP: combine pre & probably pre to pre; post and probably post to post; the rest missing
AGEMENO	Age at natural menopause	Missing if pre or had uterus removed before menopause; .=if menopausal; but unknown age, or values <20 or >67 years old	
	PLCO only has categorical age at natural menopause	.=had hysterectomy; 1=<40; 2=40-44; 3=45-49;4=50-54;5=55+; 999=menopausal, but unknown age; 999=menopausal status unknown	assign 32, 42, 47, 52, 57 for categories 1, 2, 3, 4 respectively; missing for the rest
	AARP only has categorical age at natural menopause	0=50-54; 1=<45; 2=45-49;3=55+;4=surgery; 5=medical; 6=premenopausal;7=unknown	assign 52, 40, 47, 57 for categories 0, 1, 2, 3 respectively; missing for the rest

	VITAL only has categorical agemenoc	1: <=39; 2=40-44; 3=45-49; 4=50-54; 5=55+, 996=meno status unknown; 997=reason for meno unknown; 998=non-natural menopause; 999=post-meno, age unknown	assign 32, 42, 47, 52, 57 for categories 1, 2, 3, 4, 5 respectively; missing for the rest
	CTS only has categorical agemenarchec	A=24 or less; B=35-39; C=40-43; D=44-46; E=47-49; F=50-52; G=53-55; H=56+	assign 34,37, 41.5, 45, 48, 51, 54, 57 respectively
REPRODUCTIVE HISTORY			
PREG	Number of pregnancies lasting > 6 months	Twins count as a single birth; 999=unknown	
	AARP only has categorical for live births	0: never had a child; 1=1;2=2;3=3-4;4=5-9;5=>=10;8=N/A;9=unknown	assign 0,1,2,3,7,10 for categories 0-5, and missing for the rest
	PLCO/VITAL: number of live birth	0=0; 1=1;2=2;3=3;4=4;5=5+; 999=unknown	assign 0,1,2,3,4, 5 for categories 0-5, and missing for the rest
AGEFIRSTB	Age at first birth	Age in years; .=unknown/missing	
	AARP only has categorical for age at first birth	0=never gave birth; 1=<16;2=16-19;3=20-24;4=25-29;6=30-34;7=>=40	assign 0, 16,17.5,22,27,32,37,42 respectively
DURBF	Duration of breastfeeding for all children combined	Duration in months; 999=unknown; missing if never parous	BGS, EPIC, Nurses2, Sisters, WLHS : continuous in months;
	CTS only has categorical durbfc	nulli, no live birth, <6m, 6-11, 12-23, 24-35 36-47, 48-59, 60+, bf_dkm, 999	assign 0, 3, 8.5, 17.5, 29.5, 41.5, 53.5,65, ., . respectively
	Nurse96 has pseudo-continous var created from categories	none, <1, 1-3, 4-6, 7-11, 12-17, 18-23, 24-35, 36-47, 48+ months	assign 0,2,5,9,15,21,30,42,54 respectively

CONTRACEPTION			
TUBAL	Tubal ligation	1=reported having had a tubal ligation; 0=no report of tubal ligation; .=unknown/missing	AARP & BCDDP & CLUE & IOWA & BGS & CANADA & SCHS &WLHS don't have tubal information
OCUSE	Ever use of oral contraceptives	1=reported ever using OCs; 0=reported never using OCs; 9=unknown/missing	
	AARP doesn't have this variable. So derived from OCDURc	see ocdurc categories.	assign 0 to 0; 1, 2, & 3 to 1; missing for the rest
OCDUR	Duration of OC use	Duration in years; 0=non-user; 999=unknown duration	
	PLCO only has categorical OCDURc	0=non-user;1=10+years;2=6-9 years;3=4-5 years; 4=2-3 years; 5=<1 year; 999=unknown; 9999=use unknown	assign 0, 12, 7.5, 4.5, 2.5, 0.5 to categories 0, 1, 2, 3, 4, 5, and missing to the rest
	AARP only has categorical ocdurc	0=never (or <1 year); 1=1-4 years; 2=5-9 years; 3=10+ years; 8=N/A other gender; 9=unknown	assign 0, 2.5, 7, 12 to categories 0, 1, 2, 3 respectively, and missing for the rest
	VITAl only has categorical ocdurc	0:non-user; 1=<1; 2=1-4; 3=5-9; 4=10-14; 5=15+; 998=usage status unknown; 999=user, but duration unknown	assign 0, 0.5, 2.5, 7, 12, 17.5 to categories 0-5 respectively, and missing for the rest
	CTS only has categorical ocdurc	0='0, no oc use'; 1=<1; 2=1-2; 3=3-4; 4=5-9; 5=10-14; 6=15-19; 7=20-24; 8=25+	assign 0, 0.5, 1.5, 3.5, 7, 12, 17, 22, 25 to categories 0-8 respectively
	IOWA only has categorical octme	1='<1 month';2='2-6 mon';3='7-12mon';4='13mon-2yr';5='3-5yr';6=>5yr'	assign 0.5, 4,9.5,18.5,48,90 to categories 1-6 respectively, then convert to year by diving by 12 of the above assigned values
FAMILY HISTORY OF BREAST AND/OR OVARIAN CANCER			
HXBRCA	1st degree family history of breast cancer	1=mother; sister; or	If study asked this question, then set unknown & missing

		daughter had breast cancer ; 0=no 1st degree family history	to 0
HXOVCA	1st degree family history of ovarian cancer	1=mother; sister; or daughter had ovarian cancer ; 0=no 1st degree family history; .=unknown	If study asked this question, then set unknown & missing to 0. NYU doesn't have this information
HYSTERECTOMY/OOPHORECTOMY STATUS			
HYST	Hysterectomy status	0=No; 1=Yes; .=Unknown	set unknown to missing
OOPH	Oophorectomy status	0=No; 1=Yes - one ovary removed; 2=Yes - both ovaries removed; 3=Yes - unknown ovaries removed; .=unknown	OC3 FINAL DATA ONLY KEEP OOPH=0 AND 1
	AARP has different categories	1=both removed;2=both intact;3=other surgery to ovaries; 8=not applicable-other gender; 9=unknown	assign 2, 0, 3 to categories 1-3, 9 to missing (no value 8) respectively
POST-MENOPAUSAL HORMONE USE			
PMH	Use of any type of post-menopausal hormones	0=No; 1=Yes; 9=Unknown; missing if pre;	cps2, aarp, bcddp have direct pmh
	PLCO has extra categories, derived from post_menopausal and horm_status	0=No; 1=Yes; 9=Unknown; missing if pre; 2=Possibly-current hormone user but menopausal status unknown; 3=possibly-post-menopausal but former hormone user; 4=possibly-former hormone user but menopausal status unknown	collapse 2, 3, 4 as 1, missing for unknown
DURPMH	Duration of use of any type of PMH	Duration in years; 999=unknown; missing if pre	CTS didn't provide this information
	PLCO only provided categorical durpmhc	0=never took hormones; 1=10+ years; 2=6-9 years; 3=4-5 years; 4=2-3 years; 5=<=1 year; 9=unknown; 99=possible post-	assign 0, 12, 7.5, 4.5, 2.5, 0.5 to categories 0, 1, 2, 3, 4, 5, and missing for the rest

		menopausal hormone use; 999=unknown hormone use	
	AARP provided categorical durpmh	0=nonusers; 1=<5; 2=5-<10; 3=10+; 9=unknown; 999=unknown	assign 0, 2.5, 7.5, 12 to categories 0, 1, 2, 3, and missing to the rest
	VITAL only provided categorical durpmhc	1=<1; 2=1-4; 3=5-9; 4=10-14; 5=15+; 999= unknown	assign 0.5, 2.5, 7, 12, 17.5 to categories 1,2,3,4,5 respectively; and missing for the rest
	IOWA only has categorical estrotme	1='<1 month';2='2-6 mon';3='7-12mon';4='13mon-2yr';5='3-5yr';6=>5yr'	assign 0.5, 4,9.5,18.5,48,90 to categories 1-6 respectively, then convert to year by diving by 12 of the above assigned values
EONLY	Ever use of oral estrogen only	0=No; 1=Yes; 9=Unknown; missing if pre	BCDDP & PLCO & NYU & CLUE & CTS &IOWA didn't provide this information
DUREST	Duration of use of oral estrogen only	Duration in years; 999=unknown; missing if pre	BCDDP & PLCO & NYU & CLUE & CTS &IOWA didn't provide this information
OTHER RISK FACTORS			
ALC	alcohol intake (grams/day)		
	CTS only provided categorical alc	0=None; 1=<20 gram/day; 2=>=20gram/day	assign 0, 10, 25 gram/day to categories 0, 1, 2 respectively; and missing for the rest
SMOKE	smoking status	0=never; 1=former; 2=current	

Appendix 2: Tables from the risk model project with the OCAC describing imputation variables for missing data and the log odds ratio estimates from the prediction model

Risk factors included in invasive epithelial ovarian cancer risk prediction model and distributions and covariates used in models to impute missing values for risk factors with missing values.[a]

Risk factor	Covariates included in imputation model for Risk Factor | Distribution
SNP genotypes	Site | Multinomial-Dirichlet
Family history ovarian cancer	Site | Bernoulli
Family history breast cancer	Family history ovarian cancer, site | Bernoulli
Endometriosis	Cohort, age, site | Bernoulli
Menopausal status	Alcohol, smoking status, age, site | Bernoulli
Tubal ligation	Endometriosis, education, age, cohort, site | Bernoulli
Hysterectomy	Endometriosis, tubal ligation, family history breast cancer, family history ovarian cancer, age, cohort, site | Bernoulli
Ever used MHT	Menopausal status, hysterectomy, education, age, cohort, site | Bernoulli
Type of MHT	Ever used MHT, menopausal status, hysterectomy, education, age, cohort, site | Bernoulli
Age at menarche	Age, Cohort, site | truncated Student t
Ever used OCs	Cohort, site | Bernoulli
Duration OC use	Ever used OCs, age, cohort, site | truncated Gaussian
Number of pregnancies	Hysterectomy, tubal ligation, ever used OCs, endometriosis, education, smoking, alcohol, age, cohort, site | Poisson
Age at end of last pregnancy	Number of pregnancies, age at menarche, smoking status, education, age, cohort, site | truncated Gaussian
Ever breastfed	Number of pregnancies, smoking status, education, cohort, site | Bernoulli
Duration breastfeeding	Number of pregnancies, smoking status, education, age, cohort, site | truncated Gaussian

Abbreviations: MHT, menopausal hormone therapy; OC, oral contraceptive; SNP, single nucleotide polymorphism.
[a]Left hand side variables (i.e., risk factors) may depend on any covariates given in the right hand column.

Table 4. Estimates of log odds ratios (medians) and 95% Bayesian confidence intervals for risk factors included in the invasive epithelial ovarian cancer risk prediction model containing 11 confirmed SNPs, stratified by age at diagnosis (cases) or interview/reference age (controls) and parity status[a]

Risk Factor	Age at diagnosis/Interview <50				Age at diagnosis/Interview ≥50			
	Nulliparous		Parous		Nulliparous		Parous	
	Median	95% CI	Median	95% CI	Median	95% CI	Median	95% CI
Age	**0.0331**	**(0.0064, 0.0561)**	**0.0451**	**(0.0265, 0.0643)**	-0.0117	(-0.0319, 0.0031)	**-0.0100**	**(-0.0201, -0.0035)**
Age at menarche	**-0.1515**	**(-0.2256, -0.0499)**	**-0.0780**	**(-0.1334, -0.0059)**	0.0039	(-0.0540, 0.0810)	-0.0008	(-0.0300, 0.0284)
Ever used OCs	-0.2400	(-0.6346, 0.303)	**-0.2973**	**(-0.5989, -0.0213)**	-0.3267	(-0.7350, 0.057)	-0.0377	(-0.2310, 0.1240)
Duration OC use	**-0.1339**	**(-0.1656, -0.0854)**	**-0.1585**	**(-0.2214, -0.0973)**	**-0.1094**	**(-0.1521, -0.0737)**	**-0.0531**	**(-0.0730, -0.0035)**
Number of pregnancies	N/A	---	**-0.1585**	**(-0.2214, -0.0973)**	N/A	---	**-0.0678**	**(-0.0970, -0.0380)**
Age at end of last pregnancy	N/A	---	-0.0517	(-0.1069, 0.0066)	N/A	---	-0.0280	(-0.0770, 0.0073)
Breastfeeding	N/A	---	**-0.4614**	**(-0.6803, -0.2583)**	N/A	---	-0.0420	(-0.1690, 0.0837)
Duration breastfeeding	N/A	---	-0.0075	(-0.0156, 0.0012)	N/A	---	**-0.0090**	**(-0.0150, -0.0030)**
Tubal ligation	**-1.1995**	**(-2.2391, -0.0221)**	**-0.4090**	**(-0.6310, -0.2025)**	**-0.6586**	**(-1.2867, -0.0893)**	**-0.2489**	**(-0.371, -0.1120)**
Endometriosis	**0.6690**	**(0.1737, 1.1708)**	**0.5005**	**(0.2237, 0.7712)**	0.2781	(-0.1508, 0.6791)	**0.2303**	**(0.0460, 0.4146)**
Family history breast cancer	0.4771	(-0.0261, 1.0189)	**0.2795**	**(0.0345, 0.5275)**	-0.0344	(-0.3697, 0.3183)	**0.1910**	**(0.0545, 0.3188)**
Family history ovarian cancer	**1.4634**	**(0.5564, 2.3969)**	**1.3277**	**(0.8687, 1.7933)**	**0.8250**	**(0.0761, 1.5784)**	**0.4564**	**(0.2171, 0.6963)**
Menopausal status	-0.2624	(-0.8662, 0.2738)	0.1624	(-0.1382, 0.4571)	0.0951	(-0.3473, 0.4936)	0.1087	(-0.0850, 0.2954)
MHT[b]	0.6044	(-0.0205, 1.6658)	**1.1566**	**(1.0526, 2.0932)**	-0.2309	(-0.6467, 0.2135)	-0.0758	(-0.3340, 0.1466)
Type of MHT[b]	0.1782	(-1.1274, 0.9007)	**-1.6018**	**(-2.2594, -1.0627)**	0.0558	(-0.3481, 0.5096)	-0.1037	(-0.3220, 0.1642)
Hysterectomy	-0.6791	(-1.6651, 0.3138)	**-0.9043**	**(-1.3802, -0.4562)**	-0.5627	(-1.2558, 0.1801)	-0.0080	(-0.2630, 0.2137)
MHT * Hysterectomy	0.4261	(-0.6559, 1.6354)	**-1.3288**	**(-2.3177, -0.6610)**	0.7371	(-0.186, 1.4071)	0.2306	(-0.0780, 0.6066)
Type of MHT* Hysterectomy	0.3380	(-0.3641, 1.7669)	**1.2084**	**(0.2834, 2.8248)**	-0.0419	(-0.7749, 1.0872)	-0.0103	(-0.4820, 0.3536)
rs1243180	0.1069	(-0.0503,0.2484)	0.1110	(-0.0085, 0.2307)	0.1247	(-0.0006, 0.2636)	**0.1485**	**(0.0736, 0.2262)**
rs2072590	0.1385	(-0.0120, 0.2630)	**0.1833**	**(0.0779, 0.3112)**	**0.1473**	**(0.0168, 0.2661)**	**0.1301**	**(0.0429, 0.2010)**

rs11782652	0.0652	(-0.1273, 0.2252)	0.0650	(-0.1174, 0.2114)	0.0802	(-0.1150, 0.2442)	0.0671	(-0.0934, 0.1862)
rs10088218	**-0.1739**	**(-0.3167, -0.0150)**	**-0.1916**	**(-0.327, -0.0584)**	**-0.1869**	**(-0.3504, -0.0358)**	**-0.1595**	**(-0.2593, -0.0640)**
rs757210	0.0571	(-0.0982, 0.1800)	0.0496	(-0.0822, 0.1481)	0.0353	(-0.1137, 0.1548)	**0.0968**	**(0.0282, 0.1747)**
rs9303542	**0.1733**	**(0.0352, 0.3192)**	0.1131	(-0.0277, 0.2342)	**0.1624**	**(0.0374, 0.2921)**	**0.1948**	**(0.1247, 0.2740)**
rs7651446	**0.2806**	**(0.0936, 0.4770)**	**0.2580**	**(0.0676, 0.4143)**	**0.2876**	**(0.1152, 0.4690)**	**0.2856**	**(0.1575, 0.4261)**
rs3814113	-0.0934	(-0.2231, 0.0624)	-0.1155	(-0.2244, 0.0104)	**-0.1391**	**(-0.2484, -0.0031)**	**-0.1685**	**(-0.2354, -0.0950)**
rs8170	0.0694	(-0.0688, 0.2157)	0.0289	(-0.1123, 0.1362)	0.0759	(-0.0592, 0.220)	0.0735	(-0.0080, 0.1562)
rs10069690	0.0733	(-0.0597, 0.2563)	0.0129	(-0.1194, 0.1265)	0.1240	(-0.0066, 0.2943)	**0.0864**	**(0.0118, 0.1690)**
rs12942666	**0.1042**	**(-0.0597, -0.0221)**	**0.1350**	**(0.0222, 0.2643)**	0.1216	(-0.0081, 0.2723)	0.0866	(-0.0048, 0.1726)

Abbreviations: CI, confidence interval; MHT, menopausal hormone therapy; N/A, not applicable; OC, oral contraceptive.

[a] Estimates and intervals are based on the training set only: Of the 3,370 women with age at diagnosis/interview less than 50, 649 were nulliparous, 2,697 were parous, and 24 were missing parity status. Of the 8,878 women with age at diagnosis/interview greater than or equal to 50, 1,051 were nulliparous, 7,780 were parous, and 47 were missing parity status. Missing parity status was imputed for the final model. Estimates and intervals in bold font are statistically significant at an alpha level of 0.05.

[b] MHT is coded as never use (0) and ever use (1). Type of MHT is coded as never use (0), estrogen-only MHT use (0), and all other MHT use (1). As such, among women without hysterectomy, the effect of estrogen-only MHT use is defined by MHT and the effect of all other MHT use is defined by the (MHT) + (Type of MHT). Among women a hysterectomy, the effect of estrogen-only MHT is defined by (MHT) + (MHT*Hyst) and the effect of all other MHT use is defined by (MHT) + (Type of MHT) + (MHT*Hyst) + (Type of MHT*Hyst).

Appendix 3: Results for ovarian cancer risk factor associations by histology

Associations of hormonal and reproductive factors with ovarian cancer subtypes

Exposure	Pooled cohorts[a] All cancers N= RR (95% CI)	Serous N=3384 RR (95% CI)	Endometrioid N=580 RR (95% CI)	Mucinous N=334 RR (95% CI)	Clear cell N=271 RR (95% CI)	p-diff (between histologic types)
Parity**						
Ever/never	0.70 (0.64-0.76)	0.78 (0.70-0.87)	0.47 (0.38-0.59)	0.67 (0.48-0.93)	0.34 (0.25-0.45)	**3.49^{-08}**
Number of children, continuous, per 1 child	0.98 (0.98-0.99)	0.93 (0.90-0.95)	0.79 (0.74-0.84)	0.93 (0.86-1.01)	0.68 (0.60-0.76)	**1.17^{-12}**
Number of children, categorical						
0	1.00 (ref.)	1.00 (ref.)	1.00 (ref.)	1.00 (ref.)	1.00 (ref.)	
1	0.85 (0.76-0.95)	0.87 (0.75-1.01)	0.76 (0.57-1.01)	0.72 (0.45-1.16)	0.61 (0.41-0.91)	
2	0.75 (0.68-0.82)	0.87 (0.77-0.97)	0.50 (0.39-0.64)	0.81 (0.56-1.17)	0.41 (0.30-0.57)	**4.94^{-10}**
3	0.66 (0.60-0.73)	0.77 (0.68-0.87)	0.41 (0.31-0.54)	0.64 (0.43-0.95)	0.26 (0.17-0.39)	
4+	0.58 (0.52-0.64)	0.68 (0.60-0.78)	0.33 (0.24-0.45)	0.62 (0.41-0.94)	0.14 (0.08-0.25)	
Duration of breastfeeding, per 1 year[c]	0.995 (0.989-1.002)	0.99 (0.99-1.00)	0.99 (0.97-1.01)	0.98 (0.95-1.01)	1.00 (0.98-1.02)	0.56
Oral contraceptive use						
Ever/never	0.85 (0.79-0.91)	0.79 (0.73-0.85)	0.92 (0.75-1.12)	0.98 (0.76-1.27)	0.73 (0.55-0.96)	0.19
Duration of use, continuous, per 5 year increase	0.87 (0.84-0.91)	0.85 (0.81-0.89)	0.89 (0.80-0.99)	1.01 (0.88-1.15)	0.86 (0.75-1.00)	0.13
Duration of use, categorical						
Never	1.00 (ref.)	1.00 (ref.)	1.00 (ref.)	1.00 (ref.)	1.00 (ref.)	
≤1 year	0.96 (0.86-1.08)	0.93 (0.83-1.04)	1.03 (0.77-1.36)	0.90 (0.60-1.34)	0.76 (0.51-1.14)	
>1-≤5 years	0.87 (0.79-0.96)	0.83 (0.74-0.93)	0.85 (0.66-1.11)	0.80 (0.55-1.17)	0.88 (0.61-1.25)	0.64
>5-≤10 years	0.79 (0.70-0.89)	0.71 (0.63-0.82)	0.89 (0.66-1.19)	0.88 (0.58-1.34)	0.85 (0.57-1.27)	
>10 years	0.66 (0.57-0.77)	0.60 (0.51-0.71)	0.74 (0.51-1.07)	0.99 (0.62-1.57)	0.50 (0.28-0.87)	
Tubal ligation, ever/never	0.85 (0.75-0.97)	0.94 (0.81-1.09)	0.60 (0.41-0.88)	1.08 (0.65-1.80)	0.35 (0.18-0.69)	**0.003**
Age at menarche						
Continuous, per 1 year increase	0.99 (0.97-1.01)	1.00 (0.98-1.02)	0.99 (0.93-1.04)	1.02 (0.95-1.10)	0.92 (0.85-1.00)	0.29
Categorical						
≤11	1.00 (ref.)	1.00 (ref.)	1.00 (ref.)	1.00 (ref.)	1.00 (ref.)	
12	0.92 (0.84-1.01)	0.95 (0.85-1.06)	1.01 (0.78-1.29)	1.20 (0.84-1.74)	0.78 (0.54-1.12)	
13	0.95 (0.87-1.03)	0.99 (0.89-1.09)	0.94 (0.74-1.19)	1.17 (0.83-1.65)	0.78 (0.55-1.10)	0.64
14	0.93 (0.84-1.03)	0.98 (0.87-1.10)	0.81 (0.59-1.10)	1.10 (0.72-1.66)	0.84 (0.55-1.29)	
≥15	0.90 (0.81-1.00)	0.95 (0.84-1.08)	0.93 (0.69-1.26)	1.27 (0.85-1.89)	0.57 (0.35-0.93)	
Age at menopause[c]						

Continuous, per 5 year increase	1.06 (1.02-1.11)	1.05 (1.00-1.10)	1.20 (1.05-1.37)	0.96 (0.81-1.13)	1.37 (1.14-1.65)	**0.01**
Categorical						
≤40	0.91 (0.78-1.06)	0.89 (0.74-1.08)	0.65 (0.38-1.10)	1.36 (0.80-2.32)	0.15 (0.03-0.80)	
>40-≤45	0.78 (0.69-0.90)	0.87 (0.74-1.02)	0.64 (0.42-1.00)	0.69 (0.38-1.24)	0.38 (0.17-0.87)	
>45-≤50	0.94 (0.87-1.02)	0.97 (0.88-1.08)	0.84 (0.65-1.09)	0.95 (0.68-1.32)	0.94 (0.64-1.39)	0.09
>50-≤55	1.00 (ref.)	1.00 (ref.)	1.00 (ref.)	1.00 (ref.)	1.00 (ref.)	
>55	1.04 (0.90-1.20)	1.05 (0.88-1.26)	1.20 (0.78-1.84)	1.06 (0.59-1.91)	0.94 (0.44-1.99)	
HT use[d]						
Ever/never	1.35 (1.25-1.45)	1.39 (1.28-1.51)	1.65 (1.33-2.05)	1.07 (0.79-1.44)	0.89 (0.62-1.28)	**0.01**
Duration of use, continuous, per 1 year	1.037 (1.030-1.044)	1.04 (1.03-1.05)	1.06 (1.03-1.08)	1.03 (0.99-1.07)	0.92 (0.85-0.99)	**0.0004**
Duration of use, categorical						
Never	1.00 (ref.)	1.00 (ref.)	1.00 (ref.)	1.00 (ref.)	1.00 (ref.)	
≤5 years	1.13 (1.00-1.27)	1.14 (0.99-1.30)	1.47 (1.03-2.08)	0.69 (0.41-1.16)	0.85 (0.51-1.42)	
>5 years	1.58 (1.40-1.77)	1.74 (1.54-1.97)	1.71 (1.22-2.39)	0.84 (0.51-1.38)	0.62 (0.31-1.22)	**0.002**
Hysterectomy[e]	0.98 (0.91-1.06)	1.02 (0.94-1.12)	0.86 (0.67-1.09)	0.68 (0.48-0.97)	0.57 (0.37-0.89)	**0.01**

[a]Stratified on age, birth year, and cohort (pooled analysis only), adjusted for parity, and duration of OC use
[c]parous women only
[d]post-menopausal women only
[e]Also adjusted for duration of HT use
**NYU, IA, WLHS only include parous women - these are included in all the pooled analyses of parity (i.e., ever/never, continuous parity, and categorical parity), but they are excluded from the meta-analysis of ever/never and categorical parity

<u>Associations of family history, demographic and lifestyle factors with ovarian cancer subtypes</u>

	Pooled cohorts[a]					
Exposure	**All cancers N= RR (95% CI)**	**Serous N=3384 RR (95% CI)**	**Endometrioid N=580 RR (95% CI)**	**Mucinous N=334 RR (95% CI)**	**Clear cell N=271 RR (95% CI)**	**p-diff (between histologic types)**
First degree family history of breast cancer, yes/no	1.10 (1.00-1.20)	1.11 (1.00-1.24)	1.43 (1.11-1.84)	0.80 (0.52-1.23)	0.71 (0.43-1.18)	**0.02**
First degree family history of ovarian cancer, yes/no	1.39 (1.15-1.67)	1.54 (1.25-1.90)	0.99 (0.53-1.87)	1.76 (0.86-3.57)	0.93 (0.35-2.51)	0.39
Body mass index						
Continuous, per 5kg/m^2	1.01 (0.98-1.04)	0.97 (0.93-1.01)	1.08 (0.99-1.18)	1.09 (0.97-1.22)	1.05 (0.93-1.18)	**0.04**
Categorical						
<20	1.02 (0.90-1.14)	1.07 (0.93-1.23)	0.91 (0.64-1.28)	1.28 (0.84-1.95)	0.99 (0.62-1.59)	
20-<25	1.00 (ref.)	1.00 (ref.)	1.00 (ref.)	1.00 (ref.)	1.00 (ref.)	0.07
25-<30	0.97 (0.91-1.04)	0.89 (0.82-0.97)	0.98 (0.80-1.20)	1.40 (1.09-1.81)	1.30 (0.98-1.74)	

30-<35	0.98 (0.89-1.08)	0.92 (0.82-1.03)	1.08 (0.81-1.43)	1.22 (0.82-1.82)	0.96 (0.61-1.52)	
≥35	1.15 (1.01-1.31)	1.02 (0.87-1.20)	1.33 (0.93-1.90)	1.35 (0.77-2.37)	1.28 (0.73-2.24)	
Height						
Continuous, per 0.5m	1.07 (1.05-1.10)	1.07 (1.04-1.10)	1.08 (1.01-1.15)	1.01 (0.92-1.11)	1.09 (0.98-1.20)	0.66
Categorical						
<1.60m	0.87 (0.81-0.95)	0.85 (0.77-0.93)	1.00 (0.79-1.26)	0.97 (0.71-1.32)	0.87 (0.62-1.23)	
1.60-<1.65m	1.00 (ref.)	1.00 (ref.)	1.00 (ref.)	1.00 (ref.)	1.00 (ref.)	0.27
1.65-<1.70m	1.03 (0.6-1.12)	1.05 (0.96-1.15)	0.95 (0.75-1.20)	0.88 (0.64-1.20)	0.94 (0.67-1.31)	
≥1.70m	1.14 (1.05-1.24)	1.07 (0.96-1.18)	1.30 (1.03-1.65)	1.20 (0.87-1.65)	1.22 (0.87-1.71)	
Smoking						
Ever/never	1.05 (0.99-1.11)	1.04 (0.97-1.12)	0.98 (0.83-1.17)	1.43 (1.14-1.80)	0.99 (0.77-1.27)	0.05
Continuous pack-years, per 20 pack-years	1.01 (1.00-1.02)	1.01 (0.99-1.02)	0.94 (0.86-1.03)	1.02 (0.99-1.04)	0.71 (0.55-0.92)	**0.01**
Categorical pack-years						
Never	1.00 (ref.)	1.00 (ref.)	1.00 (ref.)	1.00 (ref.)	1.00 (ref.)	
≤10	1.09 (0.99-1.20)	1.08 (0.96-1.22)	1.10 (0.83-1.44)	1.35 (0.92-1.98)	0.99 (0.67-1.46)	
>10-20	1.07 (0.94-1.22)	1.08 (0.93-1.26)	0.79 (0.53-1.19)	1.58 (1.00-2.52)	0.93 (0.55-1.57)	0.06
>20-35	0.97 (0.84-1.11)	0.99 (0.84-1.15)	0.95 (0.66-1.38)	1.46 (0.93-2.30)	0.46 (0.23-0.95)	
>35	1.01 (0.88-1.15)	1.09 (0.94-1.26)	0.89 (0.59-1.34)	1.69 (1.07-2.67)	0.41 (0.18-0.93)	

[a]Stratified on age, birth year, and cohort (pooled analysis only), adjusted for parity, and duration of OC use

Appendix 4: Preliminary results for ovarian cancer risk factor associations by tumor aggressiveness

<u>Hazard ratios and 95%CIs for key ovarian cancer risk factors for rapidly fatal tumors (death within 3 years of diagnosis) versus less aggressive disease for all tumors regardless of histologic subtype</u>

	Pooled cohorts[a]		
Exposure	**Rapidly fatal N=2104 RR (95% CI)**	**Less aggressive N=1962 RR (95% CI)**	**p-diff (between rapidly fatal vs. less aggressive)**
Parity			
Ever/never	0.67 (0.59-0.76)	0.73 (0.64-0.84)	0.37
Number of children, continuous, per 1 child	0.92 (0.90-0.95)	0.88 (0.85-0.90)	**0.02**
Number of children, categorical			
0	1.00 (ref.)	1.00 (ref.)	
1	0.79 (0.67-0.94)	0.95 (0.80-1.13)	
2	0.71 (0.62-0.82)	0.81 (0.70-0.93)	0.08
3	0.68 (0.59-0.79)	0.68 (0.59-0.79)	
4+	0.62 (0.53-0.71)	0.55 (0.47-0.65)	
Duration of breastfeeding, per 1 year[c]	0.99 (0.98-1.00)	1.00 (0.99-1.01)	0.31
Oral contraceptive use			
Ever/never	0.78 (0.71-0.87)	0.85 (0.76-0.94)	0.30
Duration of use, continuous, per 5 year increase	0.84 (0.78-0.89)	0.87 (0.81-0.93)	0.39
Duration of use, categorical			
Never	1.00 (ref.)	1.00 (ref.)	
≤1 year	0.59 (0.25-1.38)	0.76 (0.34-1.67)	
>1-≤5 years	0.45 (0.22-0.94)	0.36 (0.18-0.73)	0.62
>5-≤10 years	0.16 (0.07-0.40)	0.38 (0.17-0.86)	
>10 years	0.08 (0.03-0.24)	0.09 (0.03-0.29)	
Tubal ligation, ever/never	1.06 (0.86-1.30)	0.69 (0.57-0.84)	**0.003**
Age at menarche			
Continuous, per 1 year increase	0.99 (0.96-1.02)	0.99 (0.96-1.02)	0.96

Categorical			
≤11	1.00 (ref.)	1.00 (ref.)	
12	0.87 (0.75-1.00)	0.95 (0.83-1.09)	
13	0.87 (0.77-0.99)	0.97 (0.85-1.10)	0.24
14	0.81 (0.70-0.95)	0.97 (0.83-1.14)	
≥15	0.95 (0.81-1.10)	0.88 (0.74-1.04)	
Age at menopause[c]			
Continuous, per 5 year increase	1.06 (1.00-1.13)	1.06 (0.99-1.13)	0.98
Categorical			
≤40	1.03 (0.82-1.29)	0.89 (0.70-1.14)	
>40-≤45	0.73 (0.60-0.90)	0.83 (0.66-1.03)	
>45-≤50	1.01 (0.90-1.14)	0.91 (0.80-1.04)	0.47
>50-≤55	1.00 (ref.)	1.00 (ref.)	
>55	1.23 (1.00-1.51)	1.08 (0.85-1.37)	
HT use[d]			
Ever/never	1.16 (0.98-1.37)	1.33 (1.13-1.58)	0.24
Duration of use, continuous, per 1 year	1.03 (1.02-1.04)	1.04 (1.03-1.05)	0.32
Duration of use, categorical			
Never	1.00 (ref.)	1.00 (ref.)	
≤5 years	1.02 (0.85-1.22)	1.16 (0.96-1.39)	0.21
>5 years	1.28 (1.07-1.54)	1.61 (1.34-1.93)	
Hysterectomy[e]	0.87 (0.77-0.98)	1.00 (0.88-1.12)	0.11
First degree family history of breast cancer, yes/no	1.01 (0.87-1.17)	1.19 (1.04-1.37)	0.10
First degree family history of ovarian cancer, yes/no	1.36 (1.01-1.83)	1.69 (1.31-2.18)	0.28
Body mass index			
Continuous, per 5kg/m^2	1.05 (1.00-1.11)	0.99 (0.94-1.04)	0.06
Categorical			
<20	1.00 (ref.)	1.00 (ref.)	0.07
20-<25	1.15 (0.97-1.38)	0.93 (0.78-1.12)	

25-<30	1.01 (0.90-1.12)	0.88 (0.79-0.98)	
30-<35	1.07 (0.92-1.24)	0.98 (0.84-1.14)	
≥35	1.43 (1.18-1.74)	1.01 (0.81-1.24)	
Height			
Continuous, per 0.5m	1.08 (1.05-1.12)	1.06 (1.03-1.10)	0.44
Categorical			
<1.60m	0.86 (0.76-0.97)	0.84 (0.74-0.95)	
1.60-<1.65m	1.00 (ref.)	1.00 (ref.)	0.72
1.65-<1.70m	1.00 (0.88-1.12)	1.02 (0.90-1.15)	
≥1.70m	1.18 (1.03-1.34)	1.08 (0.95-1.23)	
Smoking			
Ever/never	1.13 (1.03-1.24)	1.02 (0.93-1.12)	0.14
Continuous pack-years, per 20 pack-years	1.06 (1.00-1.12)	0.94 (0.88-1.01)	**0.01**
Categorical pack-years			
Never	1.00 (ref.)	1.00 (ref.)	
≤10	1.11 (0.95-1.30)	1.15 (1.00-1.32)	
>10-20	1.16 (0.96-1.42)	1.07 (0.89-1.28)	0.20
>20-35	1.08 (0.89-1.31)	0.95 (0.79-1.14)	
>35	1.17 (0.98-1.41)	0.86 (0.70-1.06)	

[a]Stratified on age, birth year, and cohort (pooled analysis only), adjusted for parity, and duration of OC use

[b]p-heterogeneity between cohorts <0.01

[c]parous women only

[d]post-menopausal women only

[e]Also adjusted for duration of HT use

Hazard ratios and 95%CIs for selected ovarian cancer risk factors for rapidly fatal tumors (death within 3 years of diagnosis) versus less aggressive disease within histologic subtype (serous/poorly differentiated and endometrioid/clear cell)

	Serous only[a]			**Endometrioid and clear cell only[a]**		
Exposure	**Rapidly fatal N=1532 RR (95% CI)**	**Less aggressive N=1221 RR (95% CI)**	**p-diff (rapidly fatal vs. less aggressive)**	**Rapidly fatal N=327 RR (95% CI)**	**Less aggressive N=454 RR (95% CI)**	**p-diff (rapidly fatal vs. less aggressive)**
Parity						
Ever/never	0.72 (0.62-0.83)	0.90 (0.75-1.09)	0.06	0.49 (0.37-0.65)	0.47 (0.37-0.60)	0.85
Number of children, continuous, per 1 child	0.93 (0.90-0.96)	0.92 (0.89-0.95)	0.61	0.87 (0.80-0.94)	0.76 (0.71-0.82)	**0.01**
Tubal ligation, ever/never	1.12 (0.88-1.43)	0.84 (0.67-1.06)	0.09	1.04 (0.59-1.83)	0.49 (0.31-0.78)	0.06
Hysterectomy[e]	0.91 (0.79-1.04)	1.08 (0.93-1.25)	0.08	0.58 (0.41-0.82)	0.84 (0.64-1.10)	0.10
First degree family history of breast cancer, yes/no	1.01 (0.85-1.20)	1.14 (0.95-1.36)	0.34	0.92 (0.63-1.35)	1.37 (1.03-1.81)	0.09
First degree family history of ovarian cancer, yes/no	1.25 (0.87-1.80)	1.90 (1.39-2.58)	0.09	1.81 (0.88-3.71)	1.01 (0.50-2.04)	0.26
Body mass index						
Continuous, per 5kg/m^2	1.02 (0.96-1.08)	0.96 (0.90-1.02)	0.16	1.05 (0.92-1.19)	1.00 (0.90-1.11)	0.54
Categorical						
<20	1.00 (ref.)	1.00 (ref.)	**0.03**	1.00 (ref.)	1.00 (ref.)	0.35
20-<25	1.23 (1.01-1.51)	0.88 (0.70-1.12)		1.24 (0.82-1.90)	0.91 (0.64-1.31)	
25-<30	0.95 (0.84-1.08)	0.86 (0.75-0.99)		1.26 (0.98-1.62)	0.91 (0.72-1.13)	

30-<35	0.99 (0.83-1.18)	1.01 (0.84-1.22)		1.03 (0.68-1.56)	0.76 (0.53-1.09)	
≥35	1.32 (1.05-1.67)	0.84 (0.63-1.12)		1.45 (0.84-2.52)	1.24 (0.82-1.88)	
Smoking						
Ever/never	1.10 (0.99-1.23)	1.05 (0.93-1.18)	0.52			
Continuous pack-years, per 20 pack-years	1.09 (1.02-1.17)	0.94 (0.86-1.02)	**0.01**	0.94 (0.78-1.14)	0.88 (0.74-1.04)	0.57
Categorical pack-years						
Never	1.00 (ref.)	1.00 (ref.)	**0.03**	1.00 (ref.)	1.00 (ref.)	0.33
≤10	1.07 (0.89-1.29)	1.23 (1.02-1.47)		1.20 (0.86-1.70)	1.01 (0.76-1.33)	
>10-20	1.11 (0.88-1.41)	1.20 (0.95-1.51)		1.35 (0.88-2.07)	0.72 (0.47-1.09)	
>20-35	1.03 (0.82-1.30)	0.93 (0.73-1.19)		0.84 (0.50-1.40)	0.72 (0.48-1.08)	
>35	1.34 (1.10-1.65)	0.86 (0.66-1.12)		0.73 (0.41-1.28)	0.75 (0.48-1.16)	

[a]Stratified on age, birth year, and cohort (pooled analysis only), adjusted for parity, and duration of OC use

[b]p-heterogeneity between cohorts <0.01

[c]parous women only

[d]post-menopausal women only

[e]Also adjusted for duration of HT use

Appendix 5: Additional proposed projects using the OC3 infrastructure and the status

Project Name	Proposed by	Date	Status
NSAIDs and ovarian cancer risk	Wentzensen, Trabert	Feb. 2012	Funding obtained, analysis started
Endometriosis and ovarian cancer risk	Wentzensen, Trabert	Feb. 2012	Funding obtained, analysis started
CRP/inflammatory risk factors and ovarian cancer risk	Poole, Tworoger	Nov. 2012	R03 submission March 2015
Androgens and ovarian cancer risk	Kaaks, Fortner	Oct. 2012	Funding obtained, data collection on-going
IGFs and ovarian cancer risk	Kaaks, Fortner	Oct. 2012	Funding obtained, data collection on-going
Diabetes and ovarian cancer risk	Patel, Gapster	Nov. 2012	Obtaining funding
OncoArray (GWAS)	Wentzensen, Tworoger	June 2013	NCI U19 funding (PI: Sellers), Samples sent and ½ assayed
Pre-diagnostic biomarkers and ovarian cancer risk	Tworoger, Wentzensen	Oct. 2014	U01 submission Feb. 2015